DOCUMENT RESUME

ED 039 564 CG 005 351

AUTHOR Nighswander, James K.; And Others
TITLE A Validity Study of Self-Report and Physiological Measures of Test Anxiety.
INSTITUTION American Personnel and Guidance Association, Washington, D.C.; Southern Illinois Univ., Carbondale.
PUB DATE 23 Mar 70
NOTE 34p.; Paper presented at the American Personnel and Guidance Association Convention, New Orleans, Louisiana, March 23-26, 1970

EDRS PRICE EDRS Price MF-$0.25 HC-$1.80
DESCRIPTORS Academic Achievement, *Achievement Tests, *Anxiety, *Aptitude, Elementary School Students, *Intelligence Quotient, Performance, *Predictive Measurement, Testing Problems, Test Results

ABSTRACT

This study was an investigation into the relative predictive abilities of two types of test anxiety measures. Galvanic skin response (GSR) levels obtained during achievement testing and a self-report measure of test anxiety, the Test Anxiety Scale for Children (TASC), were used as predictors of IQ and achievement test performances of 119 fifth and sixth grade subjects. It was found that the TASC accounted for little of the variance associated with test performance. Because the correlations between the GSR measures and TASC scores were quite low or negative in direction, it was suggested that the TASC may not be measuring anxiety at all, but rather some characteristic associated with aptitude. The GSR scores when combined for all students in each of the two distinct analyses produced few significant relations with test performance. The GSR demonstrated a greater relationship with achievement than IQ performance when subjects were grouped in upper and lower categories. (Author)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED039564

Paper Presented to a Meeting of the American
Personnel & Guidance Association
March 1970

A VALIDITY STUDY OF SELF-REPORT AND PHYSICLOGICAL MEASURES OF TEST ANXIETY

James R. Nighswander
Southern Illinois University

Donald L. Beggs
Southern Illinois University

Ronald Richards
Southern Illinois University

CG005351

Anxiety has been a topic of great interest to researchers in recent years. The fact that since 1950 there have been more than 1500 studies indexed under the heading of anxiety (Spielberger, 1966), and that at least 120 different procedures have been used to infer its presence (Sarbin, 1968) testifies to the zeal with which psychologists and educators have approached this phenomenon. Several books and a great many research articles have focused primarily on anxiety among school children. Sarason, Davidson, Lighthall, Wate, & Ruebush (1960) in their book Anxiety Among Elementary School Children discuss many of the problems created by anxiety in the school setting. A more recent publication by Phillips, Martin & Meyers (1969) is a review of current research and theory of anxiety as experienced by elementary school children.

Although most studies of anxiety among children have been concerned with the existence and consequences of anxiety in general, a relatively large number of investigations have been done in the area of so-called "test anxiety," a unique form of anxiety specific to evaluative situations. Sarason and his colleagues began some rather extensive investigations of test anxiety over a decade ago and have contributed greatly to the literature through many individual studies and the Yale longitudinal study (Sarason, Hill & Zimbardo, 1964; Hill & Sarason, 1966).

The profound interest in test anxiety is certainly justifiable in view of the importance of test performance in our society. The lives of nearly every member of our culture are affected by testing. Entrance into college, job placement, promotions, ability grouping and clinical diagnoses are all, in part, determined by test performance and who can deny that in all

of these uses of test results, how one performs will in some significant way affect his future success or failure in life. For these reasons we should be knowledgeable about the various phenomena which can adversely affect test performance, and as suggested by a large number of research studies, test anxiety seems to be an important one. It has been found for example that test anxiety is related not only to test performance but also to self-concept, social class, motivation, need for approval, dogmatism, and sexual differences. Unfortunately, the myriad of research studies which have discovered these various relationships have been grossly inconsistent in their findings and inconclusive as to the antecedents, nature and consequences of test anxiety.

Several investigators have blamed the measuring instruments used for the inconsistencies which exist in the literature. The most common method of measuring anxiety is the self-report scale in which the individual reports his own unique, phenomenological experience of anxiety. According to Cattel & Schier (1958) more than a hundred such tests have been developed which claim to measure anxiety. Several of the measuring instruments designed specifically for use with children are the Children's Manifest Anxiety Scale (Castenada, McCandless & Palermo, 1956), the Test Anxiety Scale for Children (Sarason, Davidson, Lighthall, & Waite, 1958) and the School Anxiety Scale (Phillips, 1966). The Test Anxiety Scale for Children (TASC) is undoubtedly the most widely known and extensively researched instrument used in anxiety investigations with elementary school populations.

An important characteristic of the TASC, and other self-report scales as well, is that its validity is solely dependent

upon the individual's accurate introspective report of his affective experiences. Such dependence creates some unique problems for the anxiety researcher. Several excellent reviews of the difficulties encountered in the self-report measurement of anxiety are offered by Krause (1961), Lazarus (1966), and Phillips, et al (1969). A common criticism of introspective self-reports is that such indicators may not have entirely consistent referents concerning the phenomenological experience of the affect. When one person reports that he experiences anxiety he may be referring to a different state than another person. Both Krause (1961) and Lazarus (1966) consider honest reporting of affect to be particularly relevant. Individuals may respond dishonestly, or if you will, defensively, for a variety of reasons. When defensive responding occurs it is largely undetectable. Some other criticisms of self-reports are social desirability responding, which means the individual being tested answers the way he thinks the tester expects him to respond, and acquiescence or yea-saying responding. Phillips, et al (1969) contend that because of the errors inherent in self-report instruments, it is likely that from 30 to 60 percent of the variance of introspective measures of anxiety can be attributed to something other than what is generally called anxiety.

Because of the problems related to relatively error free measures of anxiety with self-report scales, researchers have sought other means of inferring the existence and extent of anxiety. Behavioral indices have been offered as a possible improvement over self-report scales (Tolman, 1958), however, a more widely accepted movement has been in the direction of the physiological measurement of affect. Although physiological

research is not a recent development, it has become increasingly more widespread due to improved instrumentation and processes to reduce and analyze psychophysiological recordings rapidly and effectively.

A physiological index commonly used is the galvanic skin response (GSR). The sensitivity of the GSR to emotional arousal and the relative ease with which it is obtained are factors which have led to its extensive application in physiological research (Edelberg & Burch, 1962; Levi, 1967). The GSR was selected for this study because it is less subject to the problems associated with self-report scales, such as defensiveness, acquiescence, and social desirability responding. These sources of error are not as likely to affect the GSR since it is difficult to manifest conscious cognitive control over the functioning of the autonomic nervous system. An additional reason for the selection of the GSR was that it made continuous data available over a relatively long period of time in which several discrete events occurred. The value of numerous measures of anxiety is obvious when one considers the many situations in which anxiety can vary.

This study was not the first to investigate test anxiety with physiological indices. Greer (1966), Chambers (1967), Lissel & Littig (1962), Getting (1966), and Raphelson (1957) have all used the GSR in the study of test anxiety. The differences between their studies and the present one are the populations used and the setting in which data was collected. In all of these investigations the research was conducted in laboratory like settings in which one subject at a time was tested. In all but Greer's (1966) study, in which he used first grade students, the population samples consisted of late adolescent or adult subjects.

The purpose of the present study was to investigate test anxiety as experienced in an achievement testing situation using fifth and sixth grade subjects. The index of anxiety was the GSR which was compared with the TASC, to determine the ability of each type of measure to predict achievement and intelligence test performance. Unlike other physiological studies of test anxiety, data was collected in the normal classroom on a number of Ss at the same time.

Method

Population

The subject population consisted of 61 male and 58 female fifth and sixth grade students from three school districts in Illinois. One school each from northern, central, and southern Illinois were used in the investigation thus giving the sample a representative geographical, racial, and socioeconomic balance within the state. Twelve subjects from each of 12 classrooms were selected for participation in the study on the basis of several criteria; (a) it was necessary to have recent IQ and achievement test scores available and (b) equal numbers were desired from each of the three IQ ranges, 89 and below, 90-110, and 111 and above. Random selection of subjects was made from those who met these criteria.

Instrumentation

The three participating schools were selected because each administered the SRA Achievement Battery in fifth and sixth grades. Since analyses consisted of comparisons with previous achievement scores it was desirable that they all come from the same battery.

A short achievement test consisting of ten-minute subtests was constructed from items contained within the SRA Achievement Battery. Arithmetic and language arts tests were used in one school where SRA Achievement scores were available for these two content areas only. For the other schools, four subtests were used. The two already mentioned were supplemented by social studies and science subtests. This instrument was developed specifically for use in the present study, only for the purpose of establishing differential GSR arousal levels elicited by the various subtests. It was neither intended nor used for the purpose of measuring achievement level.

The TASC, a self-report scale of test anxiety developed by Sarason, et al (1958), was used in comparisons with the GSR measures. A copy of this instrument may be found in the Appendix.

An E & M Instrument Co. "Physiograph Six" polygraph recorder was the apparatus used to collect GSR data. Through various modifications it was possible to obtain GSR measures on twelve subjects from each participating classroom. A silver/silver chloride active electrode and a rather large silver plated reference electrode were attached to one hand and wrist of each subject. Thirty foot long wires made it possible to extend the electrodes from the Physiograph to every position in the classroom. A sodium chloride (Sanborn-Redux) electrode paste was used as a conductive medium between the metallic electrode surfaces and the skin.

Experimental Procedure

The day before the experiment was conducted the experimenters (E_1 & E_2) introduced themselves to the students and briefly described the nature of the study. All details were not revealed,

however, Ss were informed that an achievement test would be administered. On the following day, prior to the commencement of classes, and while the children were absent from their rooms, the Physiograph was moved in and positioned at the rear of the classroom behind a large folding screen. Electrodes were extended to the desks of the Ss and the wires were taped to the floor for safety and convenience purposes. When the students entered the room the 12 experimental Ss were instructed to go wash their hands carefully. Upon return their non-dominant hands were further cleaned with alcohol and electrodes were attached. The experiment then began immediately.

The first 15 minutes of the experiment were devoted to the adaptation of the feel and novelty of the electrodes. During this time E_2 showed a non-arousing film slide while E_1 calibrated the GSR instrumentation. When adaptation was achieved the 10 minute subtests were administered consecutively and the order in which they were given was randomly determined. Continuous GSR data was obtained during each of the test administrations and for a 15 minute period of time following completion of the last test. During this post-test phase of the study another set of film slides were shown.

The entire experiment was conducted in either 1-1/4 or 1-3/4 hours depending on whether two or four subtests were administered.

Analysis of Data

The GSR recordings obtained during the experiment consisted of sustained changes in basal skin resistance from a pre-established baseline. Resistance values were converted to conductance units as recommended by Lacy & Siegel (1948). The GSR conductance

and TASC self-report measures were compared to determine their differential **abilities in predicting IQ and achievement test** performance. All statistical tests were run with a multiple linear regression analysis (Kelly, Beggs & McNeil, 1969).

Results

As mentioned in a previous section, approximately one-half of the subject sample was administered a two subtest battery consisting of arithmetic and language arts tests while the other half of the sample took a four subtest battery with additional tests covering the content areas of social studies and science. Because the treatments were different it was necessary to perform separate analyses on each group. The Ss who were administered only two tests will be referred to as coming from School A, and the Ss who were administered four subtests as coming from Schools B & C. School A consisted of a total of 57 experimental Ss and Schools B & C combined had 62 Ss.

The intercorrelation matrix found in Table I indicates the relationships among IQ and achievement test scores, TASC scores, and GSR scores obtained during various phases of the experiment, for School A.

Insert Table I about here

As expected the test scores correlated well with other test scores and GSR scores correlated well with other GSR measures. The meaningful correlations, however, are those between the TASC and test performance and between GSRs and test performance. In this matrix all coefficients over .22 are significantly different than zero at the .05 level of significance. Significant negative

correlations were found to exist between the TASC and composite achievement and language arts achievement scores. The arithmetic GSR measures failed to correlate significantly with any test scores, however, the language arts GSR level was significantly related not only with language arts test performance but also with IQ, composite and arithmetic performance. It appeared that with the School A sample, in which only two subtests were administered, that the GSR level during the language arts test was more highly related to all types of test performance than either the TASC, or other GSR levels.

Table II contains the intercorrelation matrix showing relationships between test performance, TASC, and GSR variables, for Schools B & C. There are two additional test performance and corresponding GSR variables for social studies and science content areas.

The intercorrelations between the important variables in Table II are markedly lower than those in Table I.

Insert Table II about here

The TASC was found to be significantly related to science test performance only. Two GSR scores revealed significant correlations with test scores. Social studies GSR was positively related to arithmetic achievement and post-test GSR negatively related to science test performance. In all three instances of significant correlations, the coefficients were quite low.

In both Tables I and II the correlations between the TASC scores and the various GSR levels were generally low. In those few cases in which the correlations were significant or approached significance, the relationship between the variables was negative.

These results indicate that the two measures, TASC and GSR, are not measuring the same thing.

In the initial stages of analysis it was quite discouraging to discover that for at least half the sample, GSR levels were apparently unrelated to IQ and achievement test performance. For School A Ss, GSRs during language arts tests correlated with all measures of test performance and yet in Schools B & C, only two coefficients reached significance and those were quite low, thus accounting for little variance. On a closer inspection of the data it became apparent that there was perhaps a good explanation why such results were obtained. It will be recalled that the order in which the tests were administered was randomly determined for each classroom. Random ordering was done for the purpose of good experimental procedure. It was discovered in the analysis that a phenomenon had occurred during the experiment which could render the analysis based on the original design of random ordering of treatments, meaningless. What appeared to have happened was that GSR arousal tended to increase from the first to last test rather than the expected fluctuation from test to test. Although the order in which the tests were administered did influence the initial arousal levels and overall GSR responding for all tests combined, there was a definite enchancement of arousal across tests from beginning to end. In the analyses this effect became particularly significant because the GSR scores corresponding to the test over each specific subject content area were combined. The result was that GSR levels from a test administered first in order for Ss in one classroom, were combined with GSR levels for the same test given second, third, or fourth in order for Ss in other classrooms. Since we cannot equate test time one with test

times two, three, and four due to the enhancement of arousal over test administrations, such a procedure would not provide any useful information.

In view of this finding it was decided to analyze the data according to the order in which the tests occurred. Tables III, IV, V, VI, VII, and VIII contain correlations, means and standard deviations for the tests given in the various orders.

The top part of Table III is a correlation matrix of the GSR levels for the classrooms in School A in which arithmetic and language arts test were administered first. Below the broken lines are the means associated with each of the various groups.

Insert Table III about here

Although none of the correlation coefficients were significantly different from zero it is apparent that the highest correlations were associated with the GSR level obtained during the arithmetic tests. It is revealing to note the differences between both the specific test GSR means and the overall GSR means. The first figure in order is that GSR value associated with its own particular test. Immediately below is the overall GSR mean for all tests given. In Table III the arithmetic test generated a significantly (α = .05) higher specific test GSR mean, and also a greater overall GSR mean when it was administered first in order.

Table IV contains values from School A for tests given second in order.

Insert Table IV about here

Again the correlations were non-significant but in this order of testing the higher correlations were associated with the

language arts GSR level. As with the first order, the mean specific and overall mean GSRs show that when the arithmetic test was administered first it produced greater specific and overall arousal than when language arts was administered first.

Table V is the first order of testing for Schools B & C.

Insert Table V about here

The correlation matrix at the top of the table reveals nine coefficients of correlation significantly ($\alpha = .05$) different than zero. It is noteworthy that all correlations with arithmetic GSR are negative in value while with language arts the relationship between GSR level and test performance is positive.

The specific and overall GSR means, like those of Tables III and IV, show the differential effects of the various tests on initial arousal. As in School A, Ss in Schools B & C were also most highly aroused by the arithmetic test and the arousal for the overall testing situation was significantly greater when the arithmetic was given first than when other tests were administered first.

Tables VI and VII include correlation matrices, means and standard deviations for the second and third orders of testing for Ss in Schools B & C.

Insert Tables VI & VII about here

No significant correlations between the GSR levels and test performance were found. The specific and overall GSR means are relatively meaningless since it is not indicated what specific test or tests preceded them in order.

The columns in which there are no figures indicates that the tests over subject content areas associated with those specific GSRs were not given in that order. In Table VI, for example, the arithmetic test was not administered to any subjects in the second order. Neither social studies, nor science were administered in the third order.

The fourth order of testing data is shown in Table VIII.

Insert Table VIII about here

One significant correlation, that between GSR level for science and science test performance, was found. Again the GSR means are of little meaningful value because the knowledge of preceding test order is not indicated. In the fourth order of testing the language arts test was not administered.

The last four analyses were comparisons of the TASC and overall GSR means for Ss grouped into upper and lower 25%s on composite achievement and IQ scores.

Tables IX and X are the composite achievement groups and IQ groups respectively, for School A Ss.

Insert Tables IX and X about here

The mean TASC scores for the lower 25% groups in composite achievement and IQ scored significantly (α = .05) higher on the self-report test anxiety instrument, than the upper 25% groups. The same relationship between GSR means was found for the composite achievement but not the IQ groups.

Tables XI and XII represent figures for the same analyses as found in Tables IX and X but in these cases the Ss are from

Schools B & C. Unlike School A results, the upper and lower groups in both achievement and IQ did not score significantly different on the TASC. In the composite achievement groups the GSR means were significantly ($\alpha = .05$) different with the lower group showing the higher mean score, but as before the GSR means for IQ groups were non-significant.

Discussion

A very apparent finding of the study was that correlations between the TASC and GSR measures were quite low and often negative. If we can make the assumption that anxiety is a state of organismic arousal a crucial question is raised concerning the validity of the TASC as a measure of anxiety.

Certainly the TASC has face validity because the items do **appear** to measure anxiety. The TASC also possesses predictive validity in that it can predict, with varying degrees of accuracy, IQ and achievement test performance. However, the TASC does not appear to have the kind of validity most necessary and desirable for an anxiety instrument, i.e. construct validity. There is reason to believe that the TASC is <u>not</u> measuring anxiety as anxiety is currently defined. What is the TASC measuring, then, if not anxiety? Earlier in the paper it was suggested that acquiescence and social desirability responding are sources of error in self-report instruments. Researchers (Crandall, Crandall & Katkovsky, 1965) have studied the influence of these variables on test taking behavior and they have found that a significant correlation exists between IQ and agreeing or social desirability types of responding. Low IQ individuals were more likely to answer "yes" to positively stated items than high IQ people, regardless of item content. The

TASC contains 30 items, all stated in the positive. "Yes" answers on all items are interpreted as indications of anxiety experiences. What we might be observing with the TASC and other self-report anxiety instruments is a social desirability and acquiescent response set among lower intelligence groups, a tendency which could lead to the false conclusion that this group is more highly anxious than high IQ groups.

Construct validity was not as important a concern with the GSR as with the TASC because the GSR is a commonly accepted index of autonomic arousal. There was a problem, however, in using GSR levels obtained at different times to predict test performance. The difficulty seemed to be partly a function of the unexpected phenomenon of increasing arousal from the beginning to the end of the testing situation. The enhancement of arousal over time proved to have a negative effect on the overall data analysis because it inevitably led to correlations between GSR levels and test scores which were of little interpretable value.

The most interesting and immediately practical findings of the study seemed to be related, first to the order in which tests were administered, and second to the fact that arousal tended to increase steadily from the beginning to the end of testing regardless of the order in which the test are administered.

Although the results of this study are far from conclusive with respect to either finding, further evidence along these lines could certainly have direct relevance to the particular order in which we administer achievement tests and also the amount of time allowed between tests for generated arousal to subside. The data presented here suggests that perhaps the arithmetic test within

achievement batteries should be administered later in the test order since it creates the greatest initial arousal and also the most overall arousal for the entire battery of tests.

An interesting discovery made in the study was that the highest correlations between achievement scores and GSR level were associated with the arithmetic test given in the first order. Although the GSR level was essentially uncorrelated with arithmetic test performance it was significantly related to most other achievement scores as well as IQ performance. What is even more interesting is the fact that in the SRA Achievement Battery the arithmetic subtest is administered first.

The last four analyses were included because many researchers of anxiety have dichotomized their samples into upper and lower ability groups and then compared mean anxiety scores for the two groups. Like many of the studies in which this kind of analysis was performed the results of the investigation were somewhat confusing. The TASC means were different for both IQ and achievement groups with School A Ss, and yet were not significantly different for Schools B & C Ss. The GSR level means were significantly different for achievement ability groupings in all schools and yet nonsignificant in all schools for IQ groupings. These results suggest that the physiological measure of anxiety (GSR) is more highly related to achievement test performance than ability performance. The results also argue for the physiological measure being a more appropriate measure for specific test anxiety than the TASC.

Summary

This study was an investigation into the relative predictive abilities of two types of test anxiety measures. Galvanic skin response (GSR) levels obtained during achievement testing and a self-report measure of test anxiety, the Test Anxiety Scale for Children (TASC), were used as predictors of IQ and achievement test performances of 119 fifth and sixth grade subjects. It was found that the TASC accounted for little of the variance associated with test performance. Because the correlations between the GSR measures and TASC scores were quite low or negative in direction it was suggested that the TASC may not be measuring anxiety at all, but rather some characteristic associated with aptitude.

The GSR scores when combined for all subjects in each of the two distinct analyses produced few significant relationships with test performance, however, the GSR levels appeared to be confounded by an ordering effect. There was a definite tendency for arousal to increase from the administration of the first to the last test, thus making the averaging of GSR levels obtained in different testing orders, somewhat meaningless.

In general the arithmetic test produced the most arousal compared to the other tests administered, and when the arithmetic test came first in order there was greater overall arousal for the entire battery of tests than when other tests were given first.

The GSR demonstrated a greater relationship with achievement than IQ performance when Ss were grouped in upper and lower categories. Physiological indices may be more appropriate for the measurement of test anxiety aroused during an achievement test, i.e. situation specific anxiety-than the TASC.

Although this investigation fell short of demonstrating the absolute superiority of the GSR measurement anxiety over traditional self-report scales it did raise some issues which could stimulate further research along these lines.

A replication of the present investigation with one slight alteration, could still provide the information bearing on the relative predictive abilities of GSR and TASC measures that was sought, but not found here. The change necessary in a replication study would be that of holding the order of testing constant for all subjects. As to the specific order, it is suggested that it correspond exactly to that of the achievement battery with which GSR comparisons are made. Such a procedure would eliminate the apparent confounding effect associated with different test orders.

There is a need to conduct further investigations into the differential effects of test order on overall arousal experienced during an achievement battery. Although the present study does provide some information on this topic, further investigations with larger samples are needed to clarify the issue.

References

Castaneda, A., McCandless, B.R. and Palermo, D.S. The childrens' form of the manifest anxiety scale. Child Development, 1956, 27, 317-326.

Cattell, R. B. and Schier, I. H. The nature of anxiety: a review of thirteen multivariate analyses composing 814 variables. Psychological Reports, Monograph Supplement, 1958, 5, 351-388.

Chambers, A.C. Anxiety, physiologically and psychologically measured, and its consequences on mental test performance. (Doctoral dissertation, University of Southern California) Ann Arbor, Mich.: University Microfilms, 1967, No. 68-5859.

Crandall, B.C., Crandall, B.J. & Latkovski, W. A children's social desirability questionnaire. Journal of Consulting Psychology, 1965, 29, 27-36.

Edelberg, R., and Burch, N.R. Skin resistance and galvanic skin reponse. Archives of General Psychiatry, 1962, 7, 163-169.

Greer, D.R. Test anxiety, physiological arousal and test performance of first grade children. (Doctoral dissertation, University of Missouri) Ann Arbor, Mich.: University Microfilms, 1966. No. 66-8976.

Kelly, F., Beggs, D., McNeil, K. Research design in the behavioral sciences: Multiple regression approach. Carbondale, Illinois: Southern Illinois University Press, 1969.

Krause, M.R. The measurement of transitory anxiety, Psychological Review, 1961, 68, 178-189.

Lazarus, R. S. Psychological stress and the coping process. New York: McGraw-Hill Book Co., 1966.

Levi, L. Emotional stress: psychological reactions, medical, industrial and military implications. New York: American Elvesier Publishing Co., 1967.

Oetting, E. R. Examination anxiety: prediction, physiological response and relation to scholastic performance. Journal of Educational Psychology, 1962, 53, 316-322.

Phillips, B. N. An analysis of causes of anxiety among children in school. Final Report, Project No. 2616, U.S.P.E. Cooperative Research Branch, Austin: The University of Texas, 1966.

Phillips, B. N., Martin, R. P., and Meyers, J. Anxiety and school related interventions: A selective review and synthesis of the psychological literature. New York: Board of Regents of the University of the State of New York, 1969.

Raphelson, A. C. The relationship among imaginative, direct verbal, and physiological measures of anxiety in an achievement situation. Journal of Abnormal and Social Psychology, 1957, 21, 485-490.

Sarason, S. B., Davidson, K. S., Lighthall, F. F., and Waite, R. R. A test anxiety scale for children. Child Development. 1958, 29, 105-113.

Sarason, S. B., Davidson, K. S., Lighthall, F. F., Waite R. R. and Ruebush, B. K. Anxiety in elementary school children. New York: John Wiley and Sons, 1960.

Sarason, S. B., Hill, K. T., Zimbardo, P. G. A longitudinal study of the relation of test anxiety to performance on intelligence and achievement tests. Monograph of the Society for Research in Child Development, Ser. #98, 1964, 29, No. 7.

Sarbin, T. R. Ontology recapitulates philology: The mythic nature of anxiety. American Psychologist, 1968, 23, 411-418.

Spielberger, C. D. Theory and research on anxiety. In C. D. Spielberger (Ed.) Anxiety and Behavior. New York: Academic Press, 1966, 361-398.

Tolman, E. C. A behavioristic account of the emotions. In Behavior and psychological man. Berkeley: University of California Press, 1958.

Table I

Intercorrelations of Test Performance, TASC, and GSR Levels for School A

	Variables	1	2	3	4	5	6	7	8	9
1.	IQ		.80	.75	.73	-.21	.07	-.31*	.07	-.10
2.	Comp. Ach.			.89	.95	-.31*	.09	-.30*	.10	-.08
3.	Ari. Ach.				.80	-.19	.07	-.36*	.00	-.12
4.	IA Ach.					-.38*	.11	-.24*	.12	.04
5.	TASC						.01	.11	-.19	.06
6.	Ari. GSR							.69	.38	.94
7.	IA GSR								.57	.89
8.	Post-test GSR									.50
9.	Mean GSR									

*p $\leq$.05

Abbreviation Code for All Tables

Comp. - Composite
Ari. - Arithmetic
LA - Language Arts
SS - Social Studies
Sci. - Science
Ach. - Achievement
GSR - Galvanic Skin Response
TASC - Test Anxiety Scale for Children
SD - Standard Deviation

Table II

Intercorrelations of Test Performance, TASC and GSR Levels for Schools B & C

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13
1. IQ		.70	.50	.68	.49	.50	-.18	.16	.17	-.00	.13	.14	.16
2. Comp. Ach.			.82	.86	.66	.60	-.17	.01	.12	.08	.06	.01	.08
3. Ari. Ach.				.57	.47	.38	-.17	.09	.17	.20*	.11	.08	.18
4. IA Ach.					.48	.45	-.01	.10	.17	.02	.10	.07	.13
5. SS Ach.						.51	-.16	-.04	.05	.09	-.04	-.16	.00
6. Sci. Ach.							-.23*	-.12	.01	-.01	-.17	-.22*	-.10
7. TASC								-.25*	-.34*	-.10	-.08	-.17	-.25*
8. Ari. GSR									.58	.40	.40	.67	.79
9. LA GSR										.39	.44	.57	.78
10. SS GSR											.52	.53	.72
11. Sci. GSR												.73	.78
12. Post-test GSR													.82
13. Aver. GSR													

*p < .05

Table III

Correlations Between GSR Levels and Test Performance and Means and Standard Deviations Associated With GSR Levels for the First Order of Testing Among School A Subjects

	Ari.-GSR	LA-GSR
IQ	.29	-.00
Comp. Ach.	.24	.03
Ari. Ach.	.17	-.04
LA Ach.	.33	.07
Specific Test Mean GSR	2.12	.90*
SD GSR	1.67	2.17
Overall Mean GSR	2.98	1.81
SD GSR	1.87	2.88
Mean Comp. Ach.	-5.80	18.73
SD Ach.	8.25	11.96
Mean Specific Test Ach.	-5.26	19.42
SD Ach.	5.73	15.28
Mean IQ	92.00	112.33
SD IQ	8.93	9.47

*$p < .05$

For Tables III-XII the "Mean Composite Achievement" and "Mean Specific Test Achievement" figures are mean change scores derived from subtracting grade equivalent scores from actual grade placement at the time of testing. Such a procedure was necessary because of different testing dates and also for equalizing 5th and 6th grade test scores.

Table IV

Correlations Between GSR Levels and Test Performance and Means and Standard Deviations Associated With GSR Levels for the Second Order of Testing Among School A Subjects

	Ari.-GSR	LA-GSR
IQ	-.01	.29
Comp. Ach.	.04	.28
Ari. Ach.	.00	.16
IA Ach.	.05	.28
Specific Test Mean GSR	2.72	3.85
SD GSR	3.74	2.17
Overall Mean GSR	1.81	2.98
SD GSR	2.88	1.87
Mean Comp. Ach.	18.73	-5.80
SD Ach.	11.96	8.25
Mean Specific Test Ach.	17.11	-6.26
SD Ach.	10.97	10.41
Mean IQ	112.33	92.00
SD IQ	9.47	8.93

*$p < .05$

Table V

Correlations Between GSR Levels and Test Performance and Means and Standard Deviations Associated With GSR Levels for the First Order of Testing Among Schools B & C Subjects

	Ari.-GSR	LA-GSR	SS-GSR	Sci.-GSR
IQ	-.58*	.46*	-.17	.00
Comp. Ach.	-.59*	.42*	-.27	.28
Ari. Ach.	.04	.46*	-.24	-.19
IA Ach.	-.52*	.41*	-.15	-.34
SS Ach.	-.49*	.45*	-.26	-.06
Sci. Ach.	-.16	-.03	-.34	-.21
Specific Test Mean GSR	1.36	.71	.79	.82*
SD GSR	1.08	1.14	1.17	1.21
Overall Mean GSR	2.77	1.83	1.90	1.30
SD GSR	1.37	1.57	1.37	1.20
Mean Comp. Ach.	10.36	2.05	-0.21	1.88
SD Ach.	13.94	13.10	10.21	8.45
Mean Specific Test Ach.	7.54	4.50	.00	7.00
SD Ach.	10.44	23.38	13.87	15.48
Mean IQ	109.72	105.75	108.05	100.66
SD IQ	11.37	11.94	12.45	8.02

*$p < .05$

Table VI

Correlations Between GSR Levels and Test Performance and Means and Standard Deviations Associated With GSR Levels for the Second Order of Testing Among Schools B & C Subjects

	Ari.-GSR	LA-GSR	SS-GSR	Sci.-CSR
IQ		.16	-.00	.02
Comp. Ach.		.04	.02	-.05
Ari. Ach.		.21	.15	-.12
LA Ach.		-.04	-.02	.05
SS Ach.		-.07	.10	-.16
Sci. Ach.		.03	.07	-.07
Specific Test Mean GSR		.99	2.00	1.36
SD GSR		1.02	1.62	1.17
Overall Mean GSR		1.30	2.17	1.90
SD GSR		1.20	1.57	1.37
Mean Comp. Ach.		1.88	5.00	-.21
SD Ach.		8.45	13.98	10.21
Mean Specific Test Ach.		-.55	10.29	7.36
SD Ach.		10.17	19.27	17.04
Mean IQ		100.66	107.16	108.05
SD IQ		8.02	11.90	12.45

*$p < .05$

Table VII

Correlations Between GSR Levels and Test Performance and Means and Standard Deviations Associated With GSR Levels for the Third Order of Testing Among Schools B & C Subjects

	Ari.-GSR	IA-GSR	SS-GSR	Sci.-GSR
IQ	.18	-.01		
Comp. Ach.	.09	.07		
Ari. Ach.	.20	.19		
IA Ach.	.09	.09		
SS Ach.	.22	.08		
Sci. Ach.	-.21	.00		
Specific Test Mean GSR	1.78	2.89		
SD GSR	2.04	1.86		
Overall Mean GSR	1.67	2.22		
SD GSR	1.48	1.43		
Mean Comp. Ach.	2.00	3.66		
SD Ach.	11.85	12.78		
Mean Specific Test Ach.	.68	7.93		
SD Ach.	8.83	23.93		
Mean IQ	104.17	108.66		
SD IQ	11.13	12.09		

*$p < .05$

Table VIII

Correlations Between GSR Levels and Test Performance and Means and Standard Deviations Associated With GSR Levels for the Fourth Order of Testing Among Schools B & C Subjects

	Ari.-GSR	IA-GSR	SS-GSR	Sci.-GSR
IQ	.20		-.15	.12
Comp. Ach.	.12		-.27	-.13
Ari. Ach.	-.11		-.14	-.07
IA Ach.	.25		-.28	-.05
SS Ach.	-.12		.04	-.14
Sci. Ach.	.08		-.28	-.51*
Specific Test Mean GSR	3.01		1.74	3.25
SD GSR	2.19		1.58	2.54
Overall Mean GSR	1.90		1.30	2.17
SD GSR	1.37		1.20	1.57
Mean Comp. Ach.	-.21		1.88	5.00
SD Ach.	10.21		8.45	13.98
Mean Specific Test Ach.	-.68		-3.77	11.03
SD Ach.	6.29		12.75	19.55
Mean IQ	108.05		100.66	107.16
SD IQ	12.45		8.02	11.90

*$p < .05$

Table IX

TASC and GSR Level Means and Standard Deviations For School A Subjects Grouped into Upper Middle & Lower Categories of Composite Achievement

	Lower 25%	Middle 50%	Upper 25%
Mean Comp.	-6.80	11.30	31.75
SD Comp.	6.89	5.46	6.38
Mean Overall GSR	2.78	1.94	1.79*
SD GSR	1.91	3.53	1.35
Mean TASC	13.86	10.50	8.56*
SD TASC	3.49	5.79	5.34

*$p < .05$

****

Table X

TASC and GSR Level Means and Standard Deviations For School A Subjects Grouped into Upper Middle & Lower Categories of Intelligence

	Lower 25%	Middle 50%	Upper 25%
Mean IQ	91.82	107.21	121.82
SD IQ	7.43	3.24	6.45
Mean Overall GSR	2.29	2.35	1.63
SD GSR	1.15	3.92	1.41
Mean TASC	13.17	10.60	8.82
SD TASC	4.17	6.14	4.86

*$p < .05$

Table XI

TASC and GSR Level Means and Standard Deviations For Schools B & C Subjects Grouped into Upper Middle & Lower Categories of Composite Achievement

	Lower 25%	Middle 50%	Upper 25%
Mean Comp.	-10.05	2.46	19.52
SD Comp.	3.42	7.26	9.76
Mean Overall GSR	2.38	2.08	1.68*
SD GSR	1.13	1.91	1.04
Mean TASC	16.00	14.32	15.29
SD TASC	7.38	6.66	6.60

*p $<$ .05

****

Table XII

TASC and GSR Level Means and Standard Deviations For Schools B&C Grouped into Upper Middle & Lower Categories of Intelligence

	Lower 25%	Middle 50%	Upper 25%
Mean IQ	92.17	107.65	119.52
SD IQ	4.91	7.66	4.49
Mean Overall GSR	2.13	2.05	1.84
SD GSR	1.09	1.77	1.59
Mean TASC	16.00	15.30	13.35
SD TASC	6.00	6.60	7.73

APPENDIX

STUDENT QUESTIONNAIRE
(TASC)

Yes No 1. Do you worry when the teacher says that she is going to ask you questions to find how much you know?

Yes No 2. Do you worry about being promoted. That is, passing from the ___ to the ___ grade at the end of the year?

Yes No 3. When the teacher asks you to get up in front of the class to read aloud are you afraid that you are going to make some bad mistakes?

Yes No 4. When the teacher says that she is going to call upon some boys and girls in the class to do arithmetic problems, do you hope that she will call upon someone else and not on you?

Yes No 5. Do you sometimes dream at night that you are in school and cannot answer the teacher's questions?

Yes No 6. When the teacher says that she is going to find out how much you have learned, does your heart begin to beat faster?

Yes No 7. When the teacher is teaching you about arithmetic, do you feel that other children in the class understand her better than you?

Yes No 8. When you are in bed at night, do you sometimes worry about how you are going to do in class the next day?

Yes No 9. When the teacher asks you to write on the blackboard in front of the class, does the hand you write with sometimes shake a little?

Yes No 10. When the teacher is teaching you about reading, do you feel that other children in the class understand her better than you?

Yes No 11. Do you think you worry more about school than other children?

Yes No 12. When you are at home and you are thinking about your arithmetic lesson for the next day, do you become afraid that you will get the answers wrong when the teacher calls upon you?

Yes No 13. If you are sick and miss school, do you worry that you will do more poorly in your schoolwork than other children when you return to school?

Yes No 14. Do you sometimes dream at night that other boys and girls in your class can do things you cannot do?

Yes No 15. When you are home and you are thinking about your reading lesson for the next day, do you worry that you will do poorly on the lesson?

Yes No 16. When the teacher says that she is going to find out how much you have learned, do you get a funny feeling in your stomach?

Yes No 17. If you did very poorly when the teacher called on you, would you probably feel like crying even though you would try not to cry?

Yes No 18. Do you sometimes dream at night that the teacher is angry because you do not know your lessons?

Yes No 19. Are you afraid of school tests?

Yes No 20. Do you worry a lot before you take a test?

Yes No 21. Do you worry a lot while you are taking a test?

Yes No 22. After you have taken a test do you worry about how well you did on the test?

Yes No 23. Do you sometimes dream at night that you did poorly on a test you had in school that day?

Yes No 24. When you are taking a test, does the hand you write with shake a little?

Yes No 25. When the teacher says that she is going to give the class a test, do you become afraid that you will do poorly?

Yes No 26. When you are taking a hard test, do you forget some thing you knew very well before you started taking the test?

Yes No 27. Do you wish a lot of times that you didn't worry so much about tests?

Yes No 28. When the teacher says that she is going to give the class a test, do you get a nervous or funny feeling?

Yes No 29. While you are taking a test, do you usually think you are doing poor?

Yes No 30. While you are on your way to school, do you sometimes worry that the teacher may give the class a test?